Open Bath Immersion Cooling: Density, Efficiency and Simplicity

Phillip E. Tuma

Advanced Application Development Specialist

3M Company

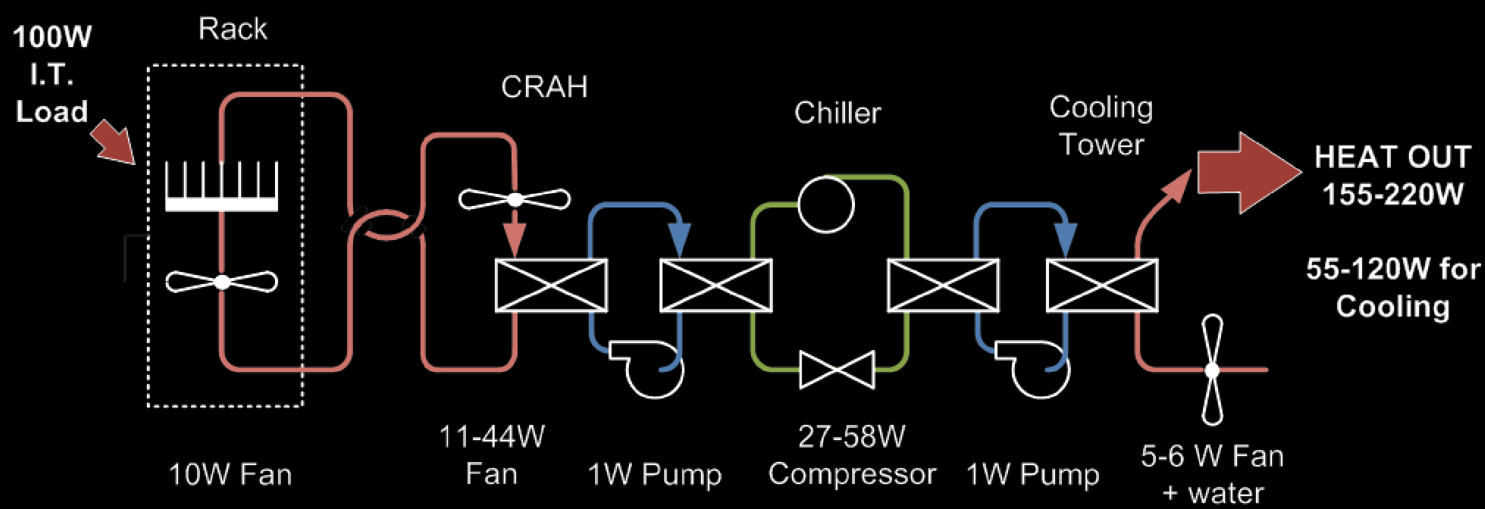
petuma@mmm.com

Outline

- Limitations of Traditional Cooling Methods
  - Traditional air, direct water/refrigerant, free air (economization)
- Open Bath Immersion Cooling Concept
  - Overview
  - Advantages
- Energy Efficiency
  - Three heat transfer processes, two power inputs
  - Actual immersion cooled server cluster
  - Energy recovery
- Power Density
  - Node level capability
  - Floorspace
- Simplicity
- Future Work
- Conclusions

Limitations of Traditional Cooling Methods

- Traditional air cooling is very inefficient and capital intensive.
  - Multiple heat transfer processes (2nd Law)
  - Mixing of hot and cold airstreams
  - Fans, pumps, blowers and compressors consume power.
  - Reliance on air as a heat transfer medium means high ΔT at heat sink.



Limitations of Traditional Cooling Methods

- Traditional water or refrigerant cooling offers efficiency and density but is expensive and complex [1].

  - Power density is greatly increased by elimination of airflow paths.
  - Energy efficiency is high because liquid and device temperatures are tightly coupled. This enables “warm” cooling techniques.
  - High temperature liquid stream permits distribution and utilization of waste heat.

  - Much of the hardware is duplicated.
  - Inability to control facility water chemistry, pressure, etc. mandates a separate water loop for the rack and an additional heat transfer process in the coolant distribution unit (CDU).
  - It is costly to design and manufacture leak-proof networks of pumps, valves, hoses, manifolds, cold plates, fittings, etc.

Water-cooled
IBM BladeCenter® HS22
IBM - Zurich

Limitations of Traditional Cooling Methods

IBM Power 775 Supercomputer Rack (>200kW)

Courtesy of International Business Machines Corporation.

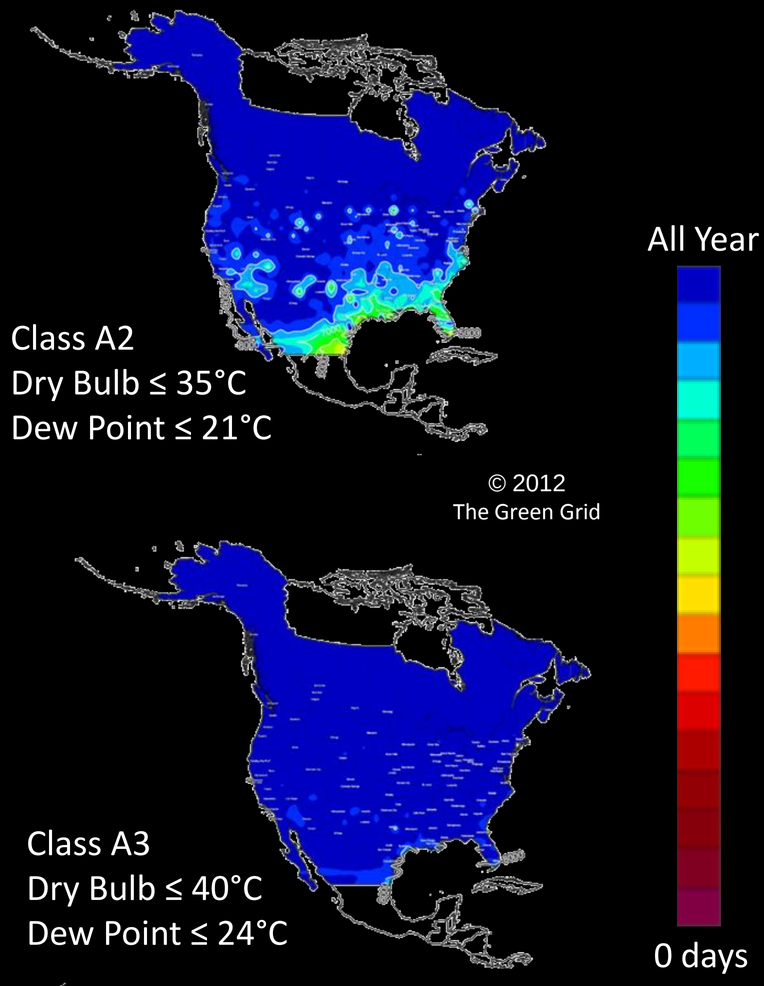
Partial list of liquid cooling hardware in one P775 rack [2,3]

Node Level	
7096	thermal interfaces by grease or gap pad
1536	DIMM heat spreaders with integrated heat pipes
1536	DIMM heat spreaders Al
264	custom bent/brazed Cu node cold plate tubes
192	CPU/Hub module cold plates
168	O-ring sealed node level tube/manifold couplings
48	water-cooled DIMM cold rails
48	node module hose sets
48	intera node PSU QDs
30	custom cast 3/4" and 1" EPDM Hose
26	custom cast QDs 3/4" bore
24	node module supply manifolds
20	high performance 120mm fans for drive bay
12	32"x21" node DIMM/VTM cold plate
12	32"x12" node hub cold plate
8	PSU cold plates, large aluminum with Cu D-tubes, and QDs
2	1"x2" L-shaped welded SS manifold assemblies, 19 ports
1	radiator core for drive bay
Rack Level	
4	90kW water conditioning unit, **$23,000 each**
1	rear door heat exchanger, **$12,000**
1	coolant fill/purge tool **$45,000**

Limitations of Traditional Cooling Methods

- "Free" air cooling is increasingly efficient but density limited.

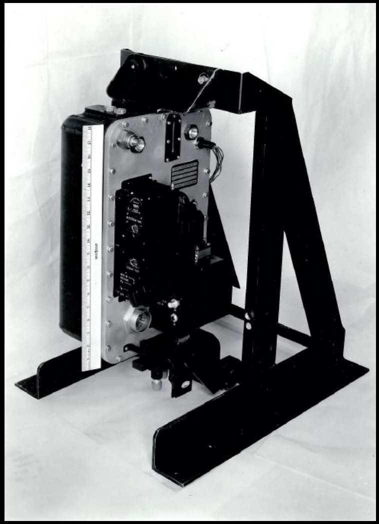
- Commodity CPUs easier to cool.
  - CPU case temperature specifications have been rising. T_{case}=85°C not unusual for 130W CPU.
  - Package and heat sink thermal performance has improved.
- Allowed Emergence of 2011 ASHRAE facility classifications A3 and A4, [4].
  - A4-compliant equipment can use air as hot as 45°C.
  - Enables free cooling just about anywhere and year round.
  - Fan power is low much of the year.
- PUEs <1.05 (excluding chassis fan power) are possible even in warm climates [5].
- Increased density comes at the expense of efficiency.



Limitations of Traditional Cooling Methods

- Traditional 2-phase immersion cooling offers efficiency and density but is expensive and complex.
  - An elegant and well-established way to capture ALL heat generated by a complex and very dense electronics assembly
  - Eliminates a complex network of plumbing.
  - It is inherently costly and complex to build hermetic electronic enclosures particularly when many conductors must penetrate that enclosure.
    - Hermetic connectors are expensive.
    - Charging/degassing complicated.
    - Servicing is complicated if one has to open a pressure vessel to access equipment.

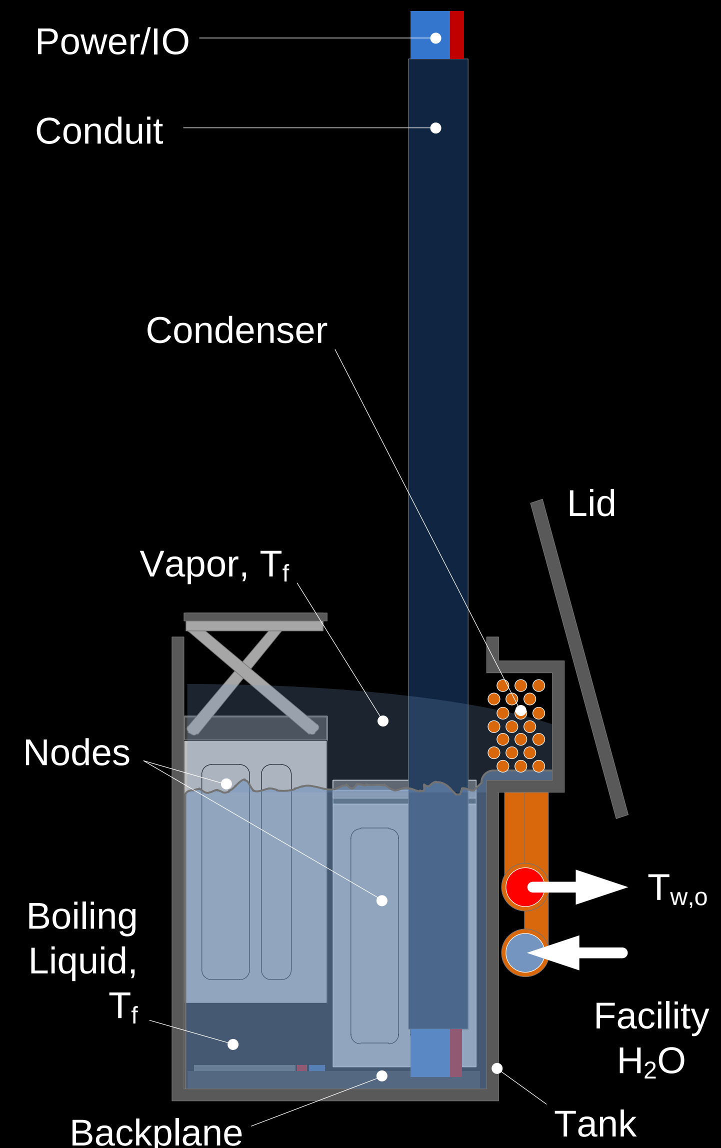
Military Enclosure, active spray (photo courtesy of Spraycool)



Airborne radar transmitter, Raytheon, -circa 1967

Open Bath Immersion Cooling Concept

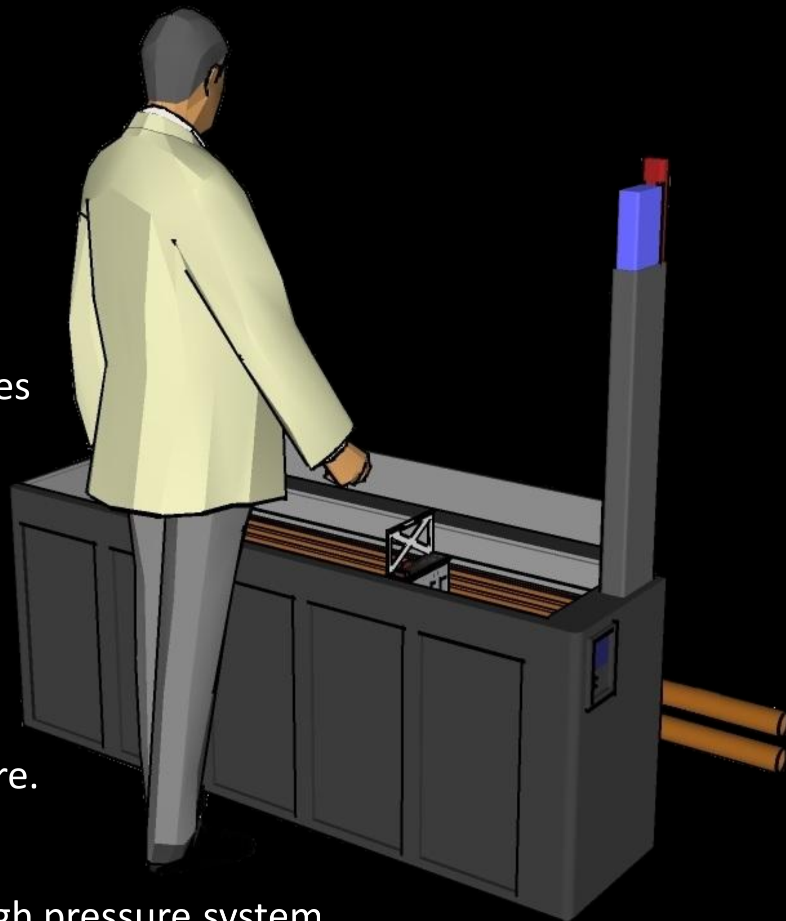
- Overview [6]
  - Servers are placed side-by-side in a lidded bath of dielectric fluid.
  - Devices cause fluid to boil.
  - Rising vapor condenses on a condenser cooled by facility water.
  - Servers plug into an immersed backplane.
  - Power and IO enter/exit through a conduit that terminates below the liquid level.
  - Servers can be hot-swapped and leave the bath dry.
  - The bath is "semi-open" as it is at atmospheric pressure. Bath breathes through a trap (not shown).



Bath Cross Section

Open Bath Immersion Cooling Concept

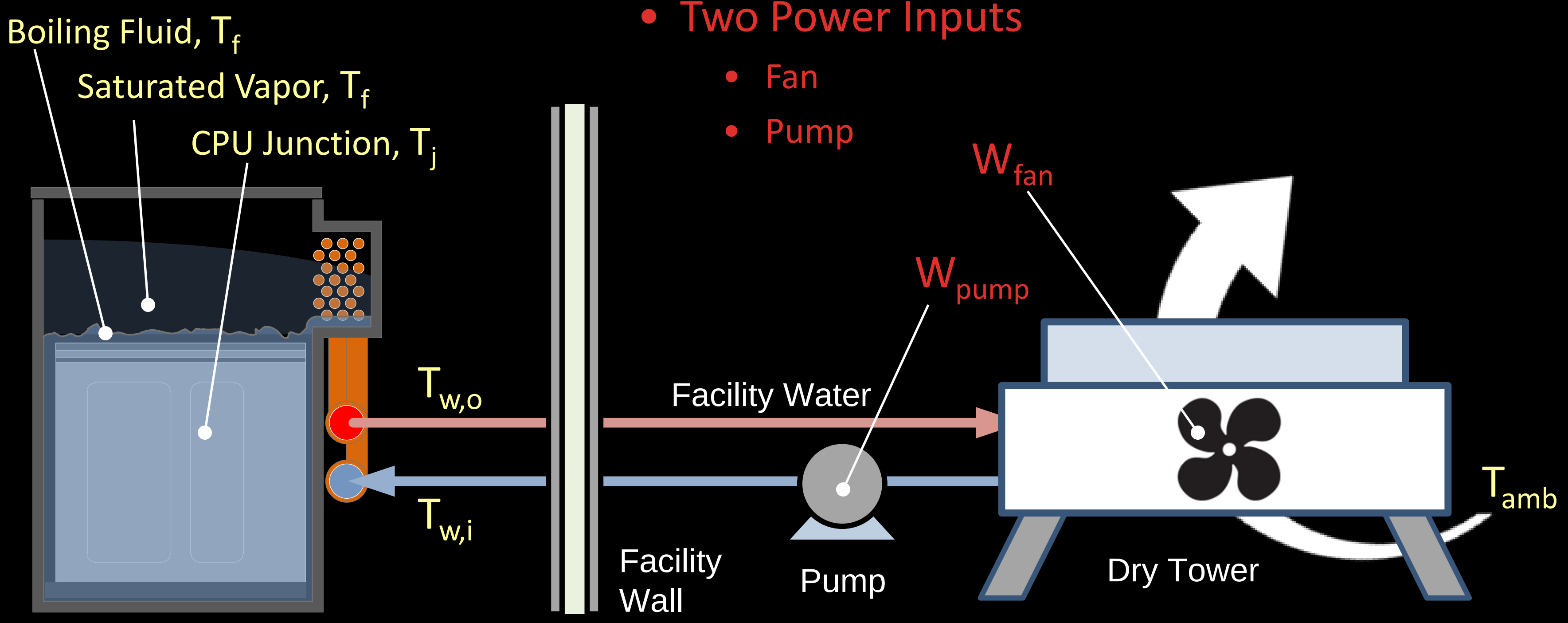
- Advantages
  - All server- and most rack-level cooling hardware are eliminated.
  - Increased packaging density can increase communication bandwidth [7]
  - Thermal efficiency is high:
    - The fluid boiling and condensation temperatures are the same. There is no advection.
    - No Secondary or ternary thermal interfaces between the silicon and fluid
    - Fluid boiled from an optimized surface that produces heat transfer coefficients >100,000 W/m^2-K.
  - All devices are kept at the same constant temperature.
  - Fluid losses occur…
    - …not at hundreds of intractable sites from a high pressure system…
    - …but at 2 sites at which the physics of fluid loss are defined and easily controlled and from a system near atmospheric pressure.



Energy Efficiency

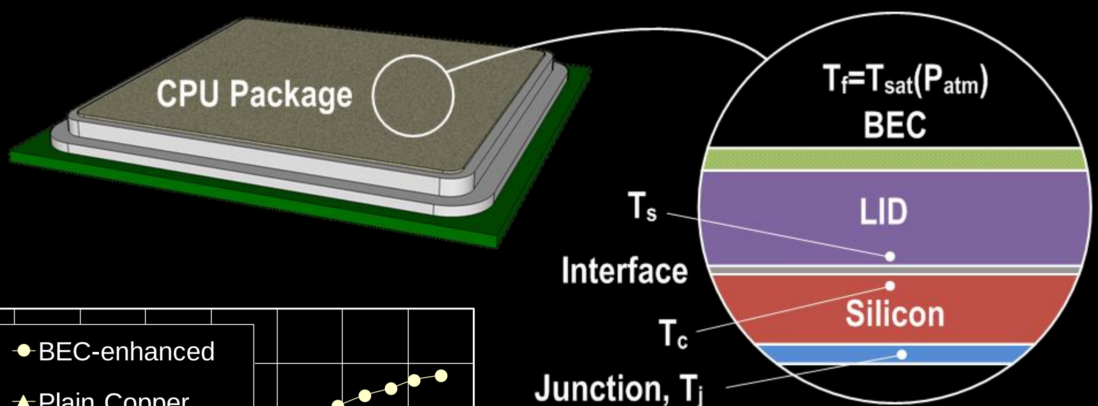
- Three Heat Transfer Processes:
  - CPU Junction-to-Fluid, ΔT_{jf} - Dictated by CPU package and boiling technology
  - Fluid-to-water, ΔT_{fw} - Dictated by condenser design/size, water flow, etc.
  - Water-to-ambient, ΔT_{wa} - Dictated by tower fan, volume, etc.

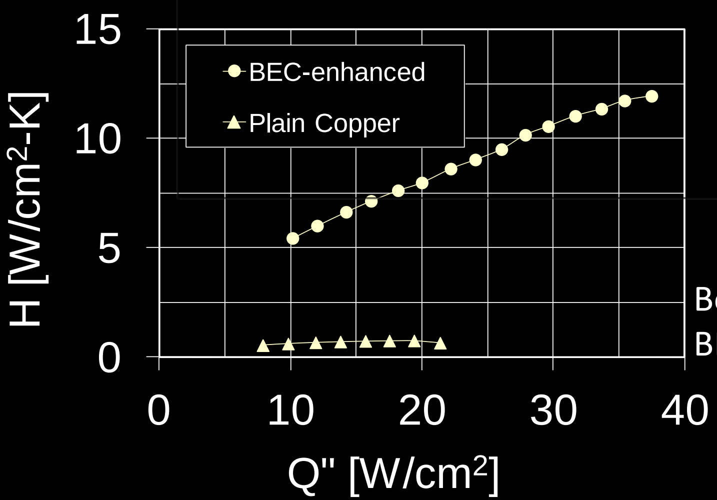
- Two Power Inputs
  - Fan
  - Pump



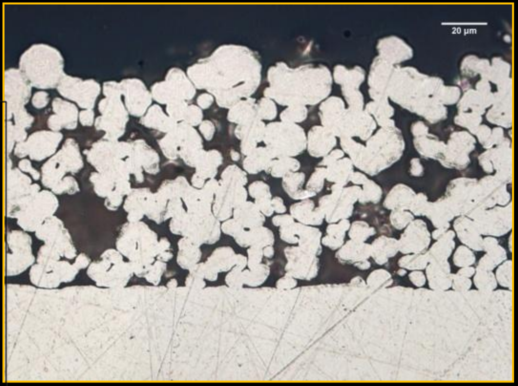
Energy Efficiency

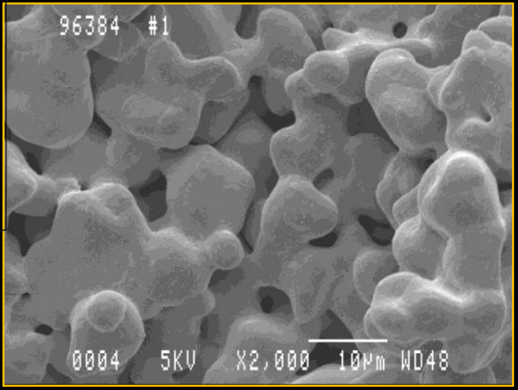
- CPU Junction-to-Fluid Performance, ΔT_{jf}
  - Boiling heat transfer enhanced with a porous copper Boiling Enhancement Coating (BEC) applied to the lid of the microprocessor.
    - 15X increase in boiling heat transfer coefficient
    - 70% increase in dryout heat flux





Boiling performance of BEC vs. plain surface





Cross section and SEM of BEC

Energy Efficiency

- Production lid has no BEC and is often too thin to adequately spread heat.
- BEC coupon can be attached with solder.
- Resultant junction-to-fluid thermal resistance, R_{jf}, depends upon
  - Chip core size, silicon thickness
  - Type of thermal interface between chip and lid

BEC coupon soldered onto AMD Processor

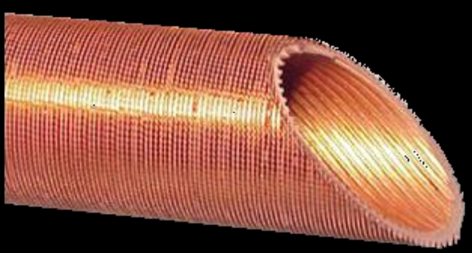
Junction-to-Fluid Thermal Resistances Measured with Real Devices (Based on "On-Chip" Sensors‡)

Device	Die Area [cm^2]	Thermal Interface	Est. Power [W]	Measured R_{jf} [°C/W]
AMD Opteron™ 83VS	2.9	solder	115	0.052
Intel I7 930†	2.6	solder	100	0.060
NVIDIA GF-100 DX11†	5.3	grease	230	0.076
IBM POWER6® dual core‡	3.6	grease	150	0.17
IBM PowerPC® 970FX	0.66	metal alloy	90	0.33

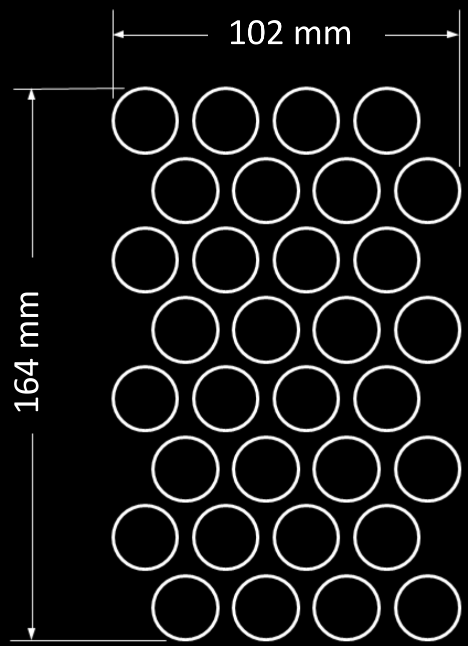
†From [8]
‡Power 6 is a modeled value [9] and the cores are ~$0.5 cm^2$

Energy Efficiency

- Fluid-to-water Performance, ΔT_{fw}
  - Best condenser design depends upon the system power.
    - Conventional radiators work well for small systems (vapor condenses on air side).
    - Enhanced tube bundles more appropriate for large scale systems (10s of kW).
  - Resultant R_{fw} depends upon design, size and water flow.



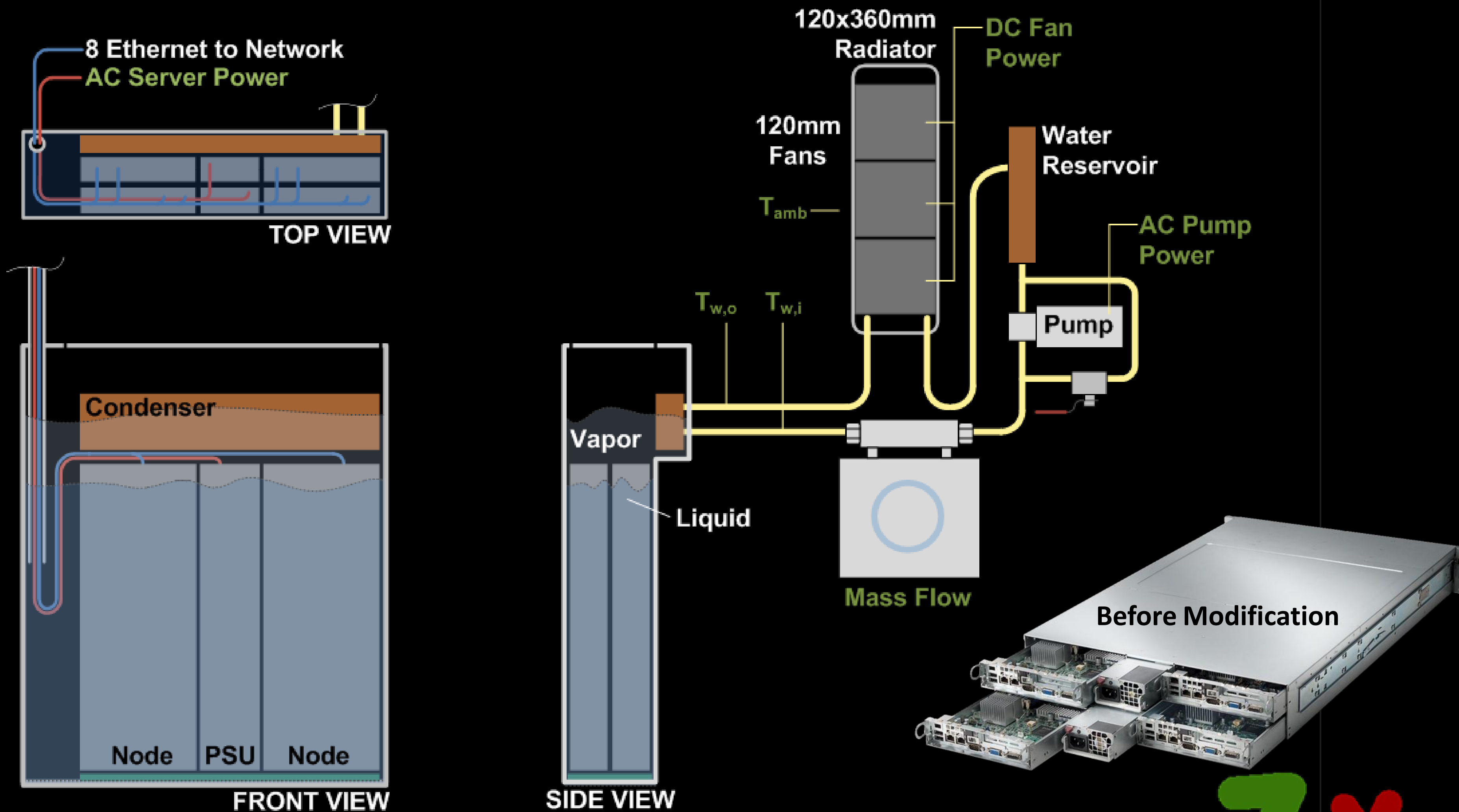
30kW condenser design based on enhanced copper tubing of the type used in large chillers. Core volume ~15 liters and log mean temperature difference is <10°C.



# Circuits	4
Passes/circuit	8
Pass L [cm]	90
Circuit L [cm]	720
Total water flow [liter/min]	60
Water Flow per circuit [liter/min]	15
Inlet ΔT [°C]	15
HTC per length [W/cm-K]	1.5
Safety Factor	0.75
Total Heat Dissipation [kW]	30.6
l/min-kW	1.96
Outlet ΔT [°C]	5.4
LMTD [°C]	9.3

- Actual Immersion-Cooled Server Cluster [10]
  - Based on 4 Supermicro MBD-H8DMT-F-B boards.
  - 8 AMD Opteron™ 83VS Processors : 115W each at full load, total power 1.2kW

Video YouTube Channel petuma1



Energy Efficiency

Simple PUE = (Fans + Pump + Server) / Server

Server Cluster Experimental Data – Fluid 1 boiling point 61°C

Control		Fluid	CPU		Air	Water			Efficiency
Fan	Pump	T_f	T_j	T_{case}	T_{amb}	$T_{w,i}$	$T_{w,o}$	liter/min-kW	PUE
Med	Med	61	67	64	22	46.4	56.1	1.47	1.008
Hi	Hi				22	34.1	38.3	3.40	1.023
Hi	Hi				40	52.1	56.3	3.40	1.023
Lo	Hi				22	52.3	56.1	check	

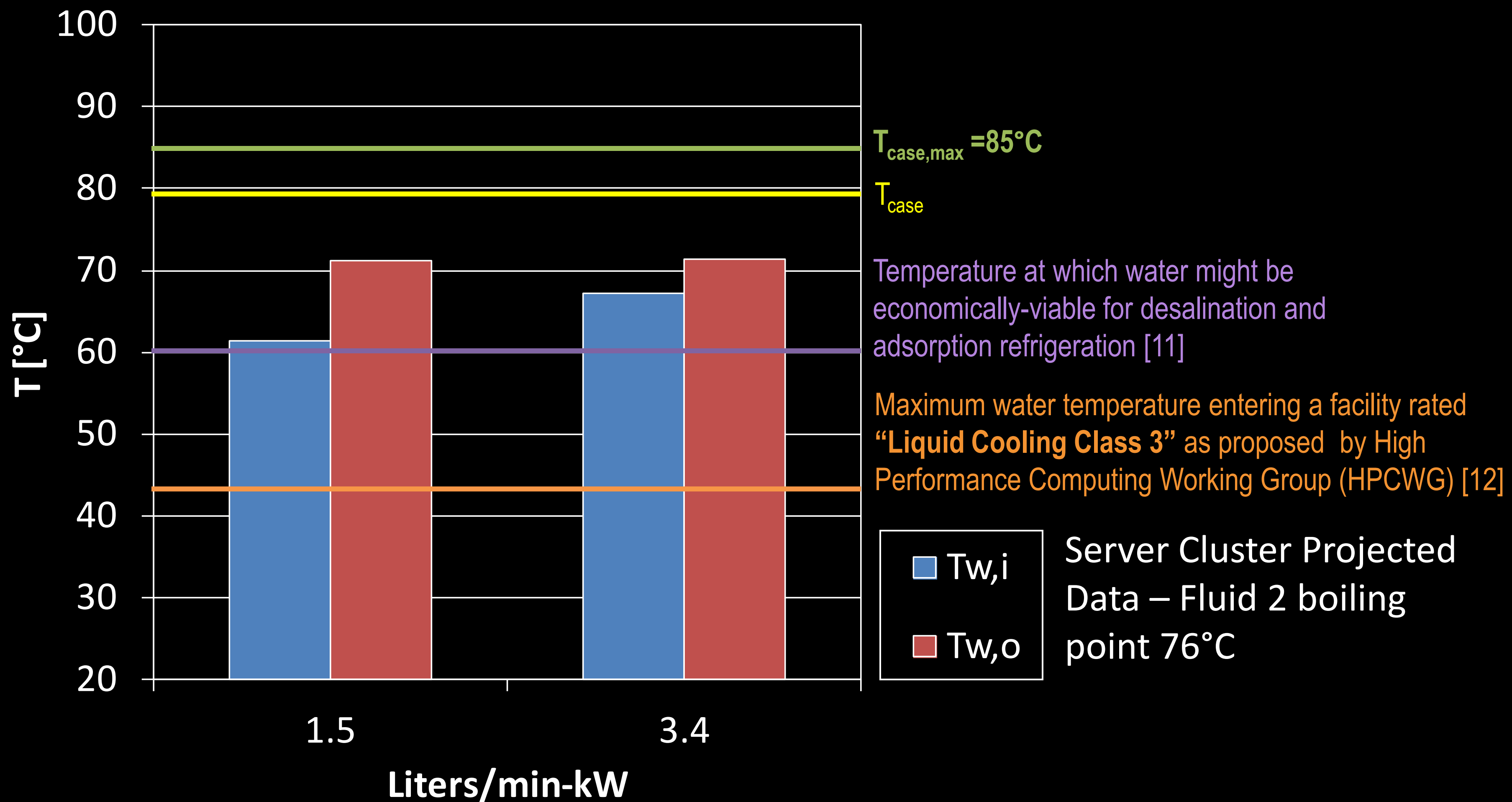
Server Cluster Projected Data – Fluid 2 boiling point 76°C

Fan	Pump	T_f	T_j	T_{case}	T_{amb}	$T_{w,i}$	$T_{w,o}$	liter/min-kW	PUE
Med	Med	76	82	79	37	61.4	71.1	1.47	1.008
Hi	Hi				55	67.1	71.3	3.40	1.023

Experimental Data
Projected

Immersion-cooled server cluster

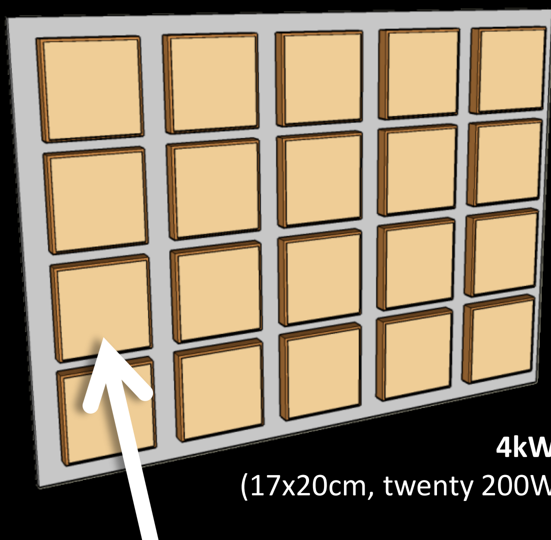
Energy Efficiency



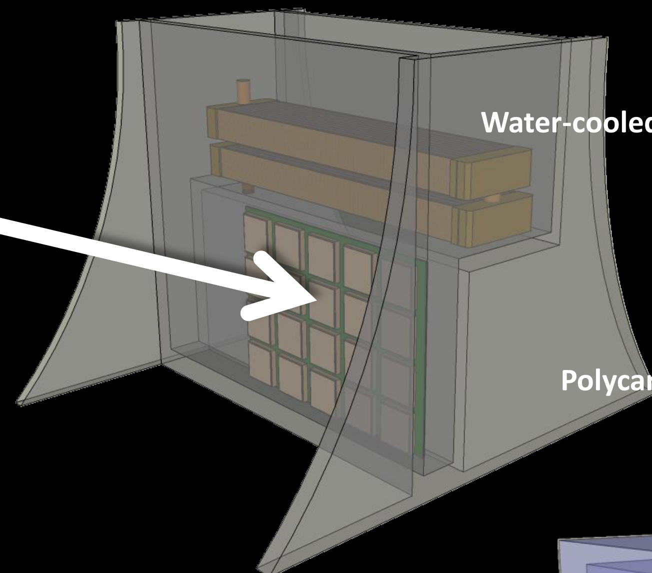
Server Cluster Projected Data – Fluid 2 boiling point 76°C

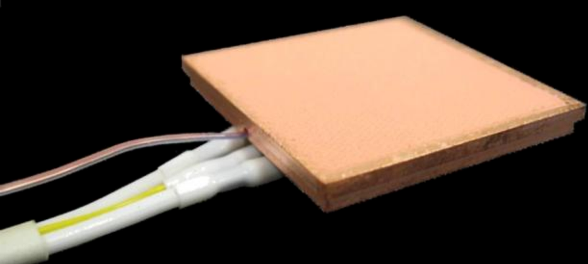
Power Density

We could not find computing hardware sufficiently dense to demonstrate the power density capability of passive 2-phase immersion. A 4 kW node simulator was therefore built.



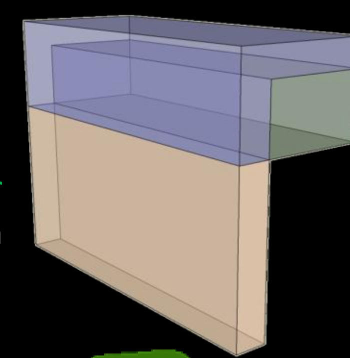
4kW blade simulator
(17x20cm, twenty 200W CPU simulators)



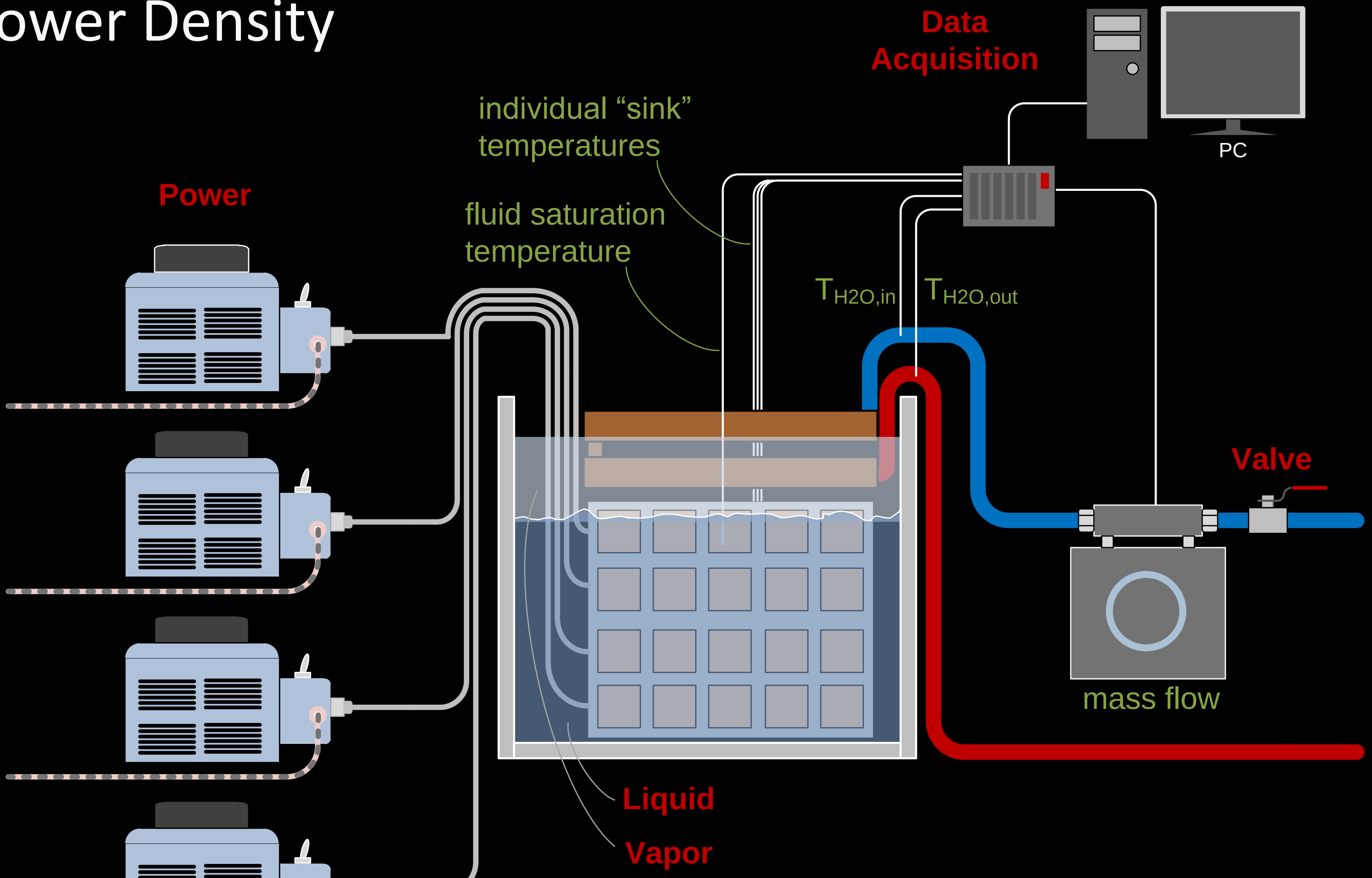


CPU Simulator Detail
(19x19mm ceramic heater,
"Sink" thermocouple in 30x30x3mm BEC-enhanced copper boiler)

Note: bath, condenser and headspace each occupy about 1 liter.

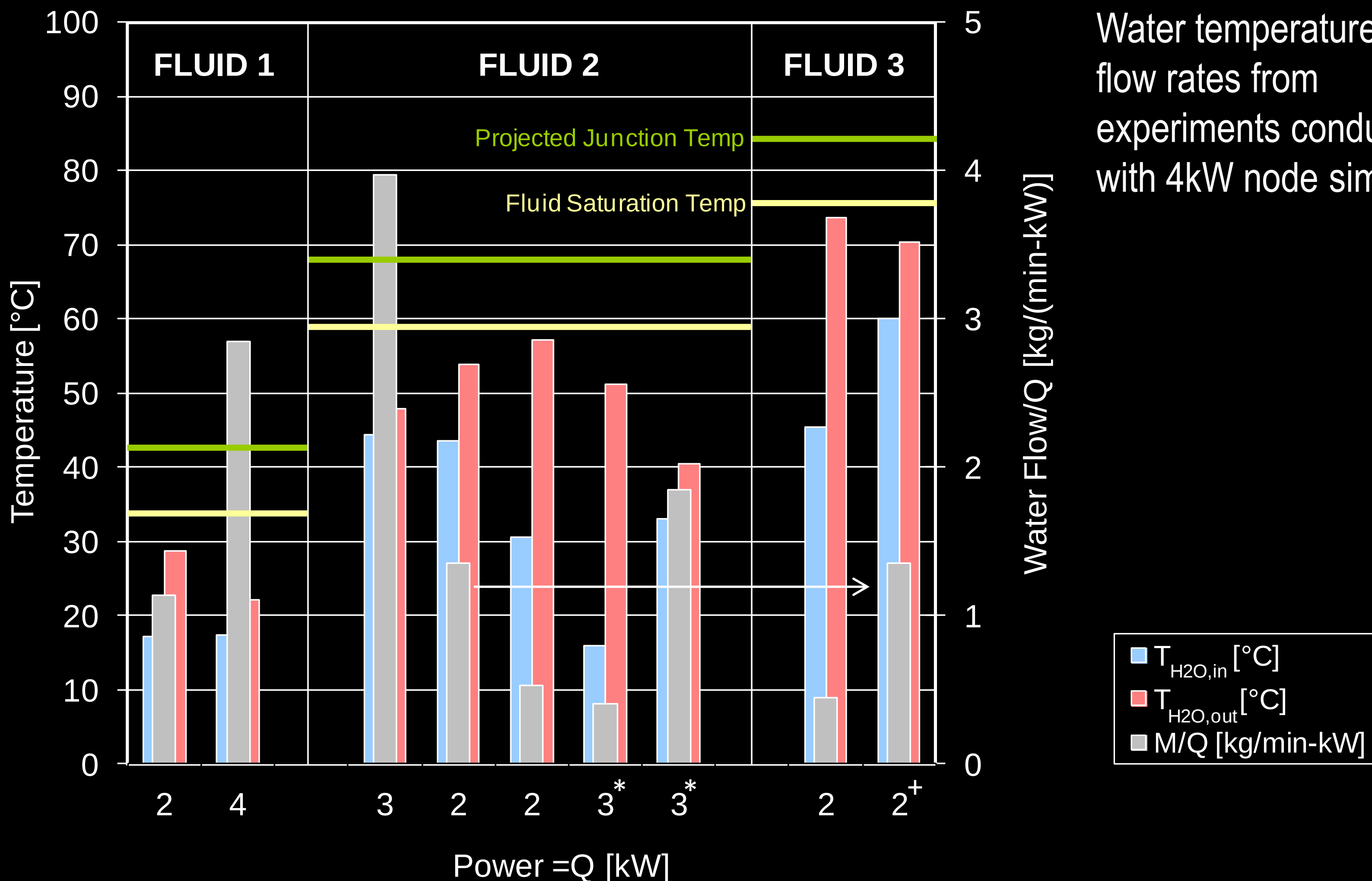


Power Density



Schematic of the test setup for experiments conducted with 4kW node simulator. 3 different fluids were used.

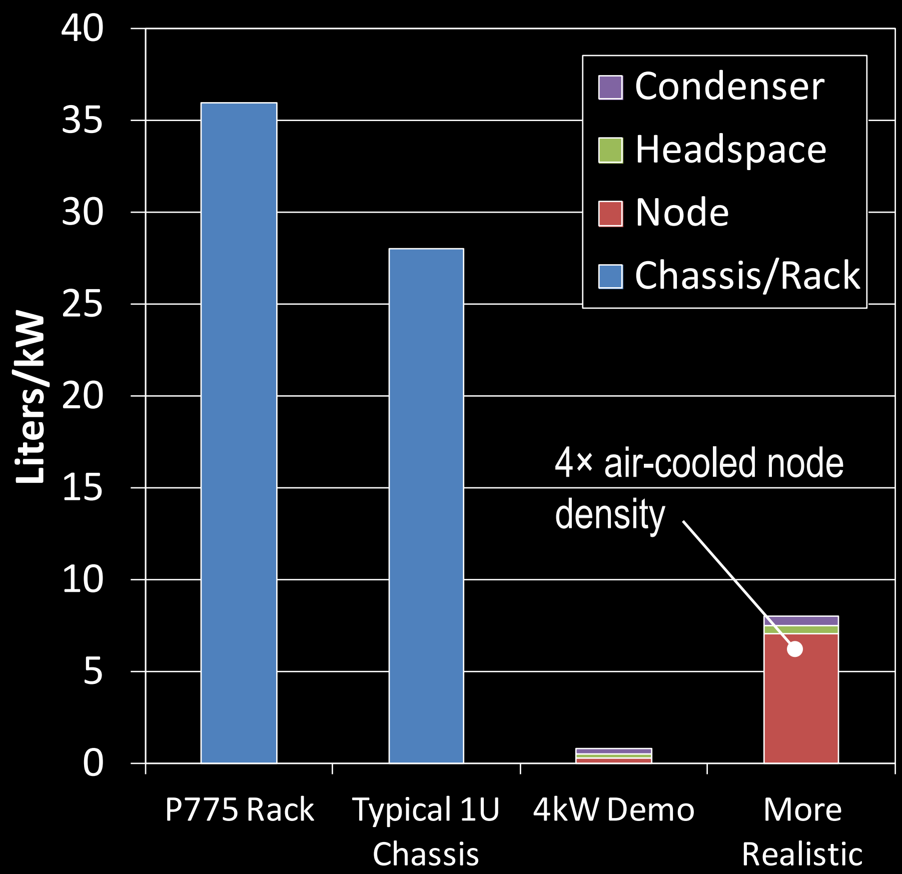
Water temperatures and flow rates from experiments conducted with 4kW node simulator.



Projected junction temperature, $T_{j,proj}$, based on a 150W CPU and average junction-to-fluid thermal resistance of 0.060°C/W.
In an optimized run, the water flow rate is adjusted so that the condenser is just able to condense the vapor being generated.
* Indicates a non-optimized run.
+ Indicates projected data – tap water would not get hot enough to perform this experiment.

Power Density

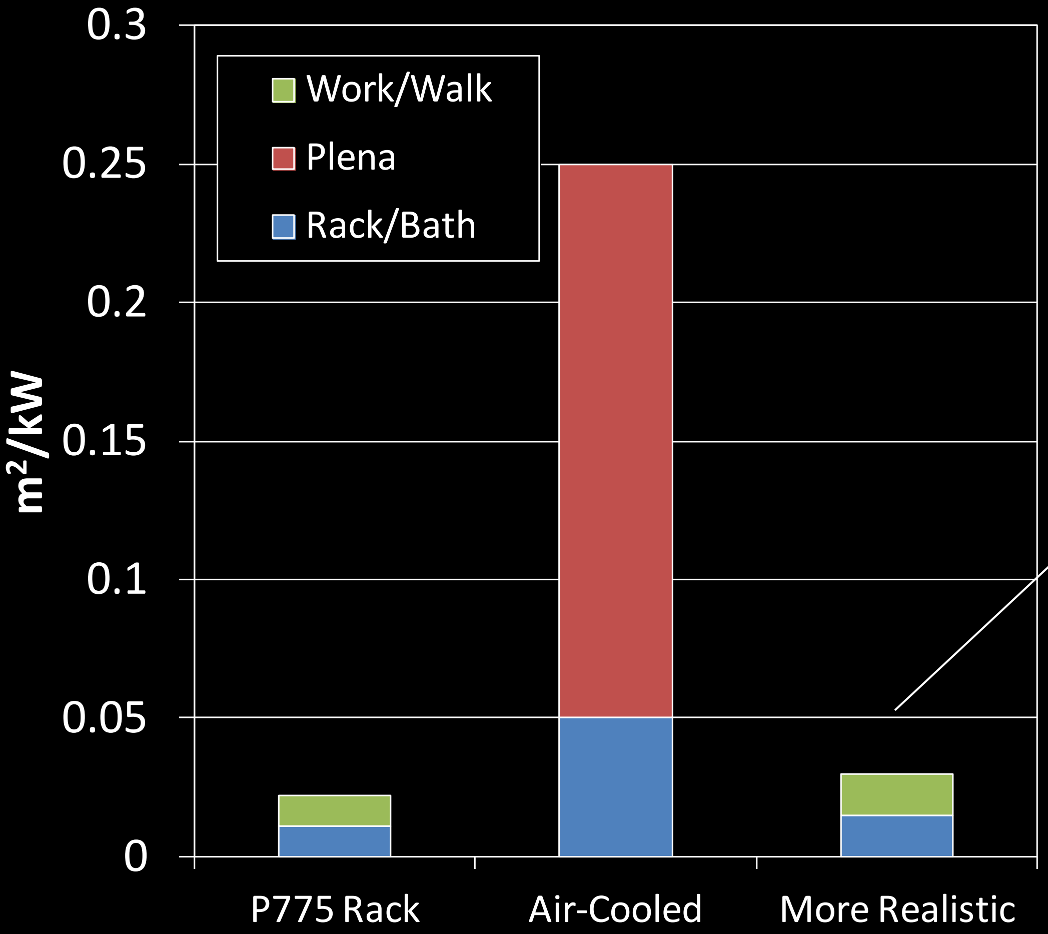
- System level capability



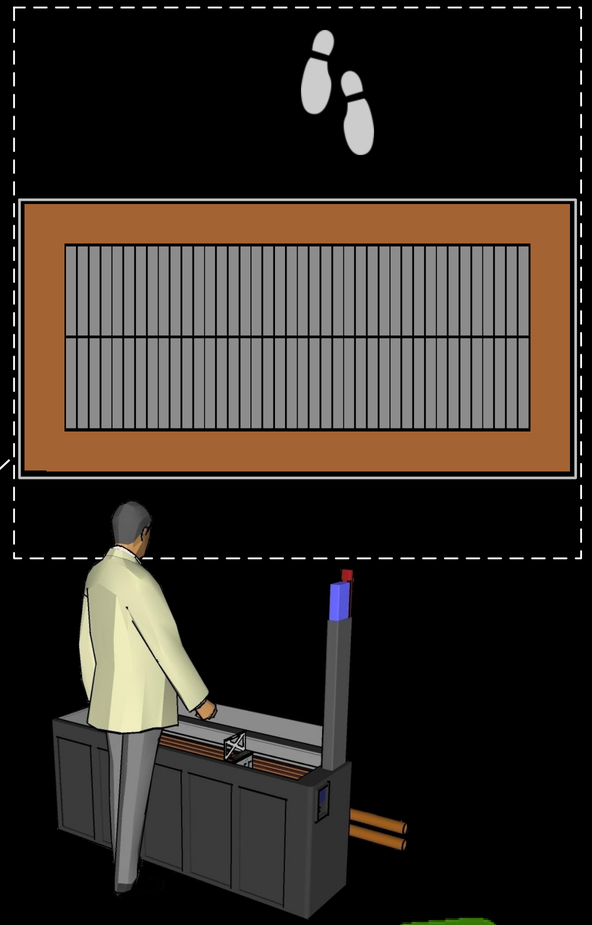
Comparison of volumetric rack and chassis power densities with system level capabilities of immersion

Power Density

- Floorspace

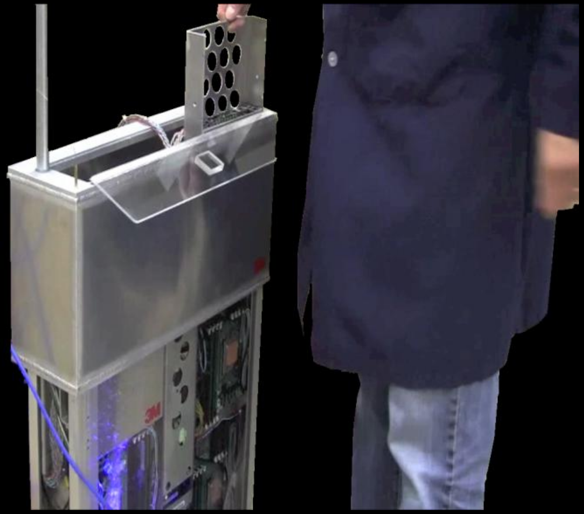


Projection: 80kW @ 4× air-cooled node density



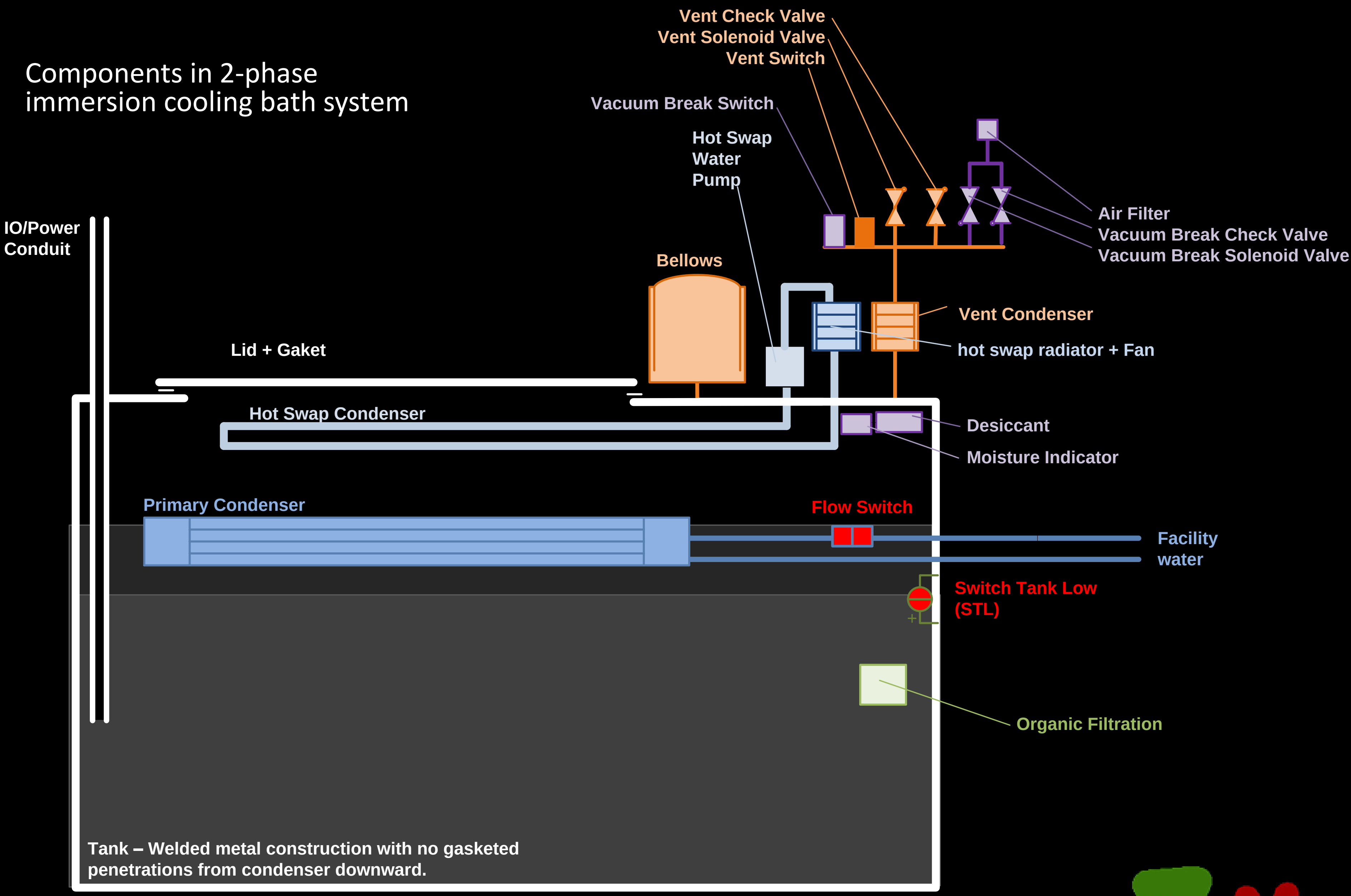
Simplicity

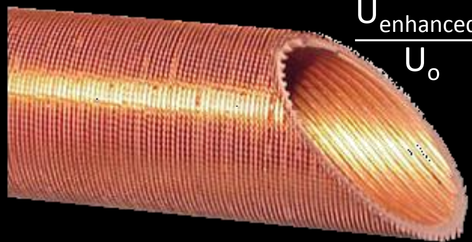
- Elimination of duplicitous node level cooling hardware
- Atmospheric pressure operation
- Passive heat transfer process – no pumps, controls, etc.
- Inexpensive subsystems functioning at bath scale:
  - Passive organic filter (activated carbon)
  - Passive desiccant to remove moisture at startup
  - Vent system
    - 2 Pressure Switches – sense ±1cm H_2O
    - Solenoid valve – vent system as needed
    - 2 Check valves – backup for solenoid
    - Bellows – Accommodate small changes in vapor height
    - Secondary condenser or "trap"
  - Hot swap condenser
    - Simple condenser active only when performing hot swaps



Hot swap experiments show very low fluid losses. These must be validated on a large scale.

Components in 2-phase immersion cooling bath system



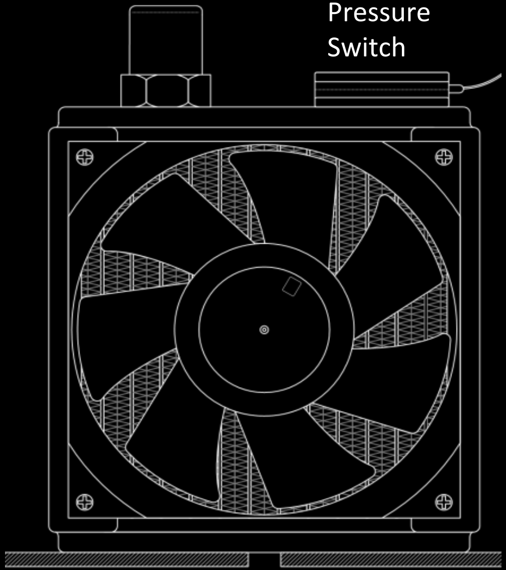


$$\frac{U_{enhanced}}{U_o} \sim 10$$

Commodity enhanced tubes like those used in low-pressure chiller condensers are ideal for use in reflux condenser.

Commodity thermostatic metering valves can modulate facility water flow to maintain a certain vapor height in the bath.

Activated carbon adsorbents will remove organic fluid contaminants that could foul the boiling surface.



Secondary condenser is active only when bellows capacity is near and system must vent air.

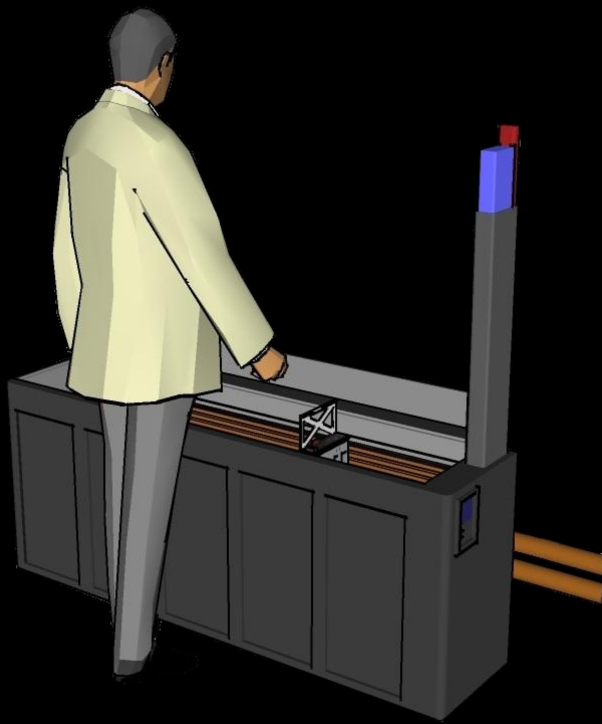
Bellows made from food packaging film can accommodate the air/vapor mixture that would otherwise be vented as the vapor zone rises during power fluctuations

Future Work

- Out of the Lab, into the Field
- Demonstrate Feasibility and Cost of Ownership
- Demonstrations pending in 3 countries
  - 75kW Immersion Bath
    - Initially heaters only,
    - To test power density and condenser performance at larger scale
  - 25kW Compute Cluster
    - 144 Processors
    - Demonstrate OBI on larger scale
  - 25kW Compute Cluster
    - Test limits for waste heat recovery
  - 25kW Compute Cluster
    - Conversion of existing super computer hardware
    - Test air-cooled equivalent alongside
  - ? kW Compute Cluster
    - Demonstrate OBI in hot, humid climate with high population density

Conclusions

- Open Bath Immersion Cooling Offers
  - Density
    - Long term node-level power density capability 4kW/liter
    - Near term floor space density similar to water-cooled supercomputers
  - Energy Efficiency
    - Potential for PUE~1.02 with 55°C ambient
    - Potential for PUE <1.01 with 37°C ambient
    - Efficient heat capture for re-use
  - Simplicity
    - Complete elimination of node level cooling hardware
    - Passive heat transfer to facility water
    - Atmospheric pressure operation
    - Simple, dry hot swappability
    - Simple controls applied not at node but bath level
    - Silent elimination of air cooling infrastructure



- Large Scale Demonstrations Pending

References

1. Ellsworth, M. and Iyengar, M., "Energy Efficiency Analyses and Comparison of Air and Water Cooled High Performance Servers," IPACK2009-89248, July 19-23, San Francisco, CA.
2. Ellsworth, M.J., "The Water-Cooled Power7 IH Supercomputing Node/System," Presentation IMAPS Advanced Thermal Workshop, Palo Alto Ca, USA, September 27-30, 2010.
3. http://www-01.ibm.com/common/ssi/rep_ca/8/897/ENUS111-058/ENUS-111-058-List_prices_2011_07_12.PDF
4. Harvey, T., Patterson, M. and Bean, J.,– "Updated Air-Side Free Cooling Maps: The Impact of ASHRAE 2011 Allowable Ranges," White Paper #46, Copyright The Green Grid, March 6, 2012.
5. "Breaking New Ground on Data Center Efficiency – How eBay's "Project Mercury" Used PUE, TCO and DCIMM Best Practices to Drive the End-to-End Data Center Ecosystem," Case Study, Copyright The Green Grid, February 2012.
6. Tuma, P.E., "The Merits of Open Bath Immersion Cooling of Datacom Equipment," Proc. 26th IEEE Semi-Therm Symposium, Santa Clara, CA, Feb. 21-25, 2010.
7. Lin, H., "Future Interconnect Demands for Super Computing," keynote address, IMAPS Topical Workshop on Advanced Interconnect Technologies, San Francisco, CA, July 14, 2010
8. Chan, B., et al, "Performance of Passive 2-Phase Immersion Cooling of Server Hardware," presentation IMAPs ATW on Thermal Management, Palo Alto, CA, USA, Sept 28-30, 2010.
9. Campbell, L. and Tuma, P., "Numerical Prediction of the Junction-to-Fluid Thermal Resistance of a 2-Phase Immersion-Cooled IBM Dual Core POWER6 Processor," IEEE Semitherm Symposium, San Jose, CA, March 18-22, 2012.
10. http://www.youtube.com/watch?v=nppkJbcKUnw
11. Brunschwiler, T., et al, "Direct Waste Heat Utilization from Liquid-Cooled Supercomputers," IHTC14-23352, Proc. 14th Int. Heat Transfer Conf., Washington D.C, August 8-13, 2010.
12. Coles, H., Ellsworth, M. and Martinez, D.J., " "Hot" for Warm Water Cooling," to be published, Int. Conf. High Performance Computing, Networking, Storage and Analysis, SC11, Nov. 12-18, Seattle, WA.

Merci